A HYBRID FORECASTING MODEL FOR GOLD PRICE PREDICTION

Dr. Pooja
PhD. Senior Lecturer in Computing Department
Westminster International University in Tashkent
Shakhzod Tulkinov
MSc. Associate Lecturer in Finance and Accounting Department
Westminster International University in Tashkent

Research area: Machine Learning, Financial Forecasting, Data Mining, Business Analytics
https://doi.org/10.5281/zenodo.21213602

Abstract. *Gold price forecasting remains a challenging task due to the nonlinear and dynamic nature of financial markets. While technical indicators are widely used to support investment decisions, their predictive performance is often limited when applied individually. This study proposes a hybrid forecasting framework that integrates traditional technical indicators with ensemble machine learning algorithms to improve the prediction of XAU/USD price movements. Historical daily gold price data spanning from July 2023 to March 2026 were collected, and a set of technical features, including the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), MACD Signal, Bollinger Bands, and calendar-based variables, were engineered. Standalone technical indicator models were first evaluated as baseline benchmarks, followed by the implementation of Random Forest, XGBoost, and Gradient Boosting models. A 5-fold time series cross-validation strategy was employed to ensure robust model evaluation while preserving the chronological order of the data. The results show that ensemble machine learning models generally outperformed traditional technical indicator approaches. Among the baseline models, Bollinger Bands achieved the highest classification accuracy of 57.53%, while the proposed hybrid model combining Random Forest and Gradient Boosting achieved the best overall performance with an accuracy of 60.14%. Feature importance and SHAP analyses identified Bollinger Band variables, particularly the lower band (BB_Low), as the most influential predictors, emphasizing the importance of volatility-related information in gold price forecasting. The proposed framework demonstrates that combining technical indicators with ensemble machine learning techniques can improve directional prediction accuracy and provide a practical decision-support tool for financial market participants.*

Keywords: *Gold Price Forecasting; XAU/USD; Machine Learning; Random Forest; Gradient Boosting; Technical Indicators; Bollinger Bands; Time Series Cross-Validation.*

1. Introduction

1.1 Background of the Study

Historically, Gold (XAU/USD) has long been used as a medium of exchange, a unit of account, and a store of value, holding a unique position in the global financial system as both a monetary asset and a commodity (Baur and Lucey, 2010; Baur and McDermott, 2010). In the contemporary world, gold is considered vital to individual investment strategies and national reserves, remaining its status as a reserve asset across centuries. Gold is often considered as a safe haven asset whose value remains or even increases, specifically during the periods of heightened economic uncertainty and geopolitical unrest, while the value of other riskier assets decreases.

Stocks and gold have negative correlation, meaning that when stock prices decline in most cases, gold price moves in the opposite direction. Gold serves as an effective hedge against inflation and systemic financial risks as it is not directly tied to the performance or stability of any single economy unlike fiat currencies, which are sensitive to inflationary pressures and monetary policy shifts (Ghosh et al., 2004; Beckmann, Berger and Czudaj, 2015). In the modern financial world, the development of algorithmic trading and advanced data analytics has resulted in an increased reliance on technical analysis. Technical indicators, including the Relative Strength Index (RSI), Bollinger Bands and Moving Average Convergence Divergence (MACD), apply historical data to identify price patterns and forecast trends. Even though those indicators are widely used in trading, when used individually, the forecasting power of those indicators is often limited because they often struggle to account for the non-linear noise and unexpected changes in the gold market.

1.2 Problem Statement

Historical price patterns repeat in predictable cycles, which is the fundamental assumption of the technical analysis. Nevertheless, typical indicators including Moving Averages, RSI and Bollinger Bands often make errors in forecasting trends because they highly rely on historical average values, which frequently fail to account for rapid changes in the gold market. When the market is highly volatile, those indicators often give conflicting signals, creating a significant issue for the analysts. For example, when the RSI shows an overbought condition, a Bollinger Bands squeeze might simultaneously indicate a forthcoming breakout, leaving the analyst in a difficult position. The main weakness of these empirical methods is their linear nature. There are complicated, non-linear relations between various market dimensions which cannot be accounted for with a simple mathematical formula. There is an increasing gap between those traditional methods and the capabilities of modern computational intelligence while a financial analyst might struggle to synthesize multiple technical signals simultaneously. When used alone, those indicators give a useful baseline. However, they do not have the multidimensional processing power to detect the hidden, non-linear relationships that identify gold price patterns. In this study, we address this gap through the development of a hybrid forecasting framework, moving beyond isolated, rule-based trading. This research explores whether algorithmic processing can overcome weaknesses of traditional analysis via ensemble machine learning architectures, Random Forest, XGBoost and Gradient Boosting in particular, integrating different technical signals as input variables. The aim is to find out whether those advanced models can successfully de-noise market data and provide better forecasting results in contrast to traditional methods.

1.3 Research Objectives

The main goal of this research is to develop a hybrid model which improves gold price forecasting. The main objectives are as follows:

- To compute and analyze key technical indicators (Bollinger Bands, RSI, MACD, and Moving Averages) for the XAU/USD pair.
- To develop baseline models using these individual technical indicators to establish a performance benchmark.
- To build advanced machine learning models, specifically Random Forest, XGBoost and Gradient Boosting, using the computed technical indicators as input features.

- To design a hybrid forecasting framework that optimally combines multiple indicators with machine learning techniques.
- To evaluate and compare model performance using statistical metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Directional Accuracy.
- To assess whether the proposed hybrid models significantly outperform traditional standalone indicator-based approaches.

1.4 Research Questions

The following research questions are answered to achieve those aforementioned objectives:

- How effective are traditional technical indicators in predicting XAU/USD price movements when used as standalone tools?
- Can machine learning models, such as Random Forest, XGBoost and Gradient Boosting, improve prediction accuracy when using these technical indicators as their primary inputs?
- Does a hybrid approach, integrating multiple indicators into a single ML framework, outperform standalone indicator models?
- Which forecasting framework demonstrated the strongest predictive performance for gold forecasting?

1.5 Significance of the Study

This research focuses on how these indicators can be most effectively integrated within a unified computational framework rather than finding the most optimal standalone indicator to use.

In this way, we move away from the conflicting debate of which single indicator is the best and we find the synergy between man-made technical indicators and machine-made pattern recognition. Moreover, this study provides a brand-new alternative to traditional econometric models through the use of ensemble learning models including XGBoost, Gradient Boosting and Random Forest, as they are efficient and capable of handling the non-linear noise of tabular financial data. The findings of the research are beneficial to institutional investors and retail traders. Finally, this study develops a methodology to improve gold price forecasting by addressing the gap between traditional technical analysis and modern algorithmic forecasting.

2. Literature Review

2.1 Conceptual Framework

2.1.1 The Relative Strength Index (RSI)

The Relative Strength Index (RSI) is one of the most widely applied momentum oscillators in financial forecasting. Introduced by Wilder (1978), RSI measures the speed and change of price movements, oscillating between 0 and 100. Numerous empirical studies have demonstrated its predictive power. For instance, Neely et al. (2014) found that RSI, when combined with other technical indicators, improved the accuracy of equity premium forecasts. Similarly, a study by Chong and Ng (2008) on the London Stock Exchange demonstrated that RSI, in conjunction with moving averages, yielded statistically significant returns. However, despite its popularity, RSI is not without limitations. Gupta (2025) noted that RSI can generate false signals in highly volatile markets, as it relies on historical price momentum, potentially missing abrupt regime changes. Despite these drawbacks, the RSI remains a cornerstone of momentum-based technical analysis, often serving as a baseline in hybrid forecasting frameworks (Sezer et al., 2020).

According to Fidelity Investments (2026), traditionally, the value of RSI above 70 means that an asset is considered overbought, giving a signal of price downward trend, and when it is below 30, it is considered as oversold, giving a signal of price upward trend as it can be seen in the following diagram.

Diagram 1. Relative Strength Index (RSI)

(Fidelity Investments, 2026)

2.1.2 The Moving Average Convergence Divergence (MACD)

The Moving Average Convergence Divergence (MACD) is a trend-following momentum indicator developed by Gerald Appel in the late 1970s. It measures the relationship between two moving averages of asset prices, typically the 12-day and 26-day exponential moving averages. When the MACD line crosses above its signal line, it is interpreted as a bullish signal; when it crosses below, it signals bearish momentum. Several studies have examined MACD’s efficacy in financial markets. For example, Lo, Mamaysky, and Wang (2000) argued that MACD, when combined with other technical indicators, could generate statistically significant returns, particularly in trend-based markets. Moreover, a study by Chong and Ng (2008) demonstrated that MACD performed well in predicting price trends on the London Stock Exchange. However, MACD is also known to lag behind sudden market reversals, as highlighted by Gupta (2025). He noted that MACD often struggles in sideways or range-bound markets, where the indicator generates frequent false signals. Despite these limitations, MACD remains a key component of many technical trading strategies, especially when combined with other indicators and machine learning models, which can enhance its predictive capacity (Fidelity Investments, 2026). Fidelity Investments (2026) claim that the MACD value exceeding 0 is considered bullish whereas the value crossing below 0 is considered bearish as it can be seen in diagram 2.

Diagram 2. The Moving Average Convergence and Divergence (MACD)

Fidelity Investments (2026)

2.1.3 Bollinger Bands

Bollinger Bands, introduced by John Bollinger in the 1980s, are a popular volatility indicator used in financial forecasting. The bands consist of a simple moving average (typically 20 periods) and two standard deviation lines above and below the moving average, forming a dynamic envelope. According to Sezer, Gudelek, and Ozbayoglu (2020), Bollinger Bands are broadly applied in assessing market volatility and identifying overbought or oversold conditions.

A study by Chong and Ng (2008) found that Bollinger Bands performed robustly in predicting price reversals, particularly in equity markets. Moreover, Neely et al. (2014) demonstrated that Bollinger Bands, when combined with other indicators, improved directional predictions of asset prices. However, Bollinger Bands also have limitations; Gupta (2025) pointed out that they can yield false signals in sideways markets when price movements oscillate between the bands without a clear trend. Despite these drawbacks, Bollinger Bands remain integral to technical analysis, especially when used alongside other indicators and machine learning methods, which capitalize on their ability to highlight volatility-driven trading signals (Fidelity Investments, 2026). Fidelity Investments (2026) claim that when the price reaches upper or lower bands, there is a likelihood of a sharp price movement in either direction. They also claim that when the price moves out of the band, there is a strong trend continuation as you can see in the following diagram.

Diagram 3. Bollinger Bands indicator

(Fidelity Investments, 2026)

2.1.4 Random Forest Machine Learning Model

Random Forest, introduced by Breiman (2001), is an ensemble learning method based on constructing multiple decision trees and aggregating their predictions. In financial forecasting, Random Forest has gained traction due to its flexibility and ability to model complex nonlinear relationships. For example, a study by Kim, Kim, and Kim (2016) applied Random Forest to predict stock market returns and found it outperformed linear models. Similarly, Chen et al. (2019) used Random Forest to forecast cryptocurrency prices, showing strong predictive accuracy in volatile markets. Despite its strengths, Random Forest can be prone to overfitting, especially when the number of trees is not adequately tuned (Gupta, 2025). Nevertheless, its interpretability, ability to handle large feature sets, and robustness in various market conditions have made it a cornerstone in financial machine learning applications (Fidelity Investments, 2026).

2.1.5 XGBoost Machine learning model

Extreme Gradient Boosting (XGBoost) is an advanced ensemble learning algorithm developed by Chen and Guestrin (2016) that builds upon the gradient boosting framework. The model constructs decision trees sequentially, with each new tree correcting the errors of the previous trees, resulting in improved predictive performance. XGBoost has gained significant popularity in financial forecasting due to its ability to handle complex non-linear relationships, missing values, and large datasets efficiently. Studies have demonstrated its effectiveness in predicting financial market movements. For instance, Zhang, Aggarwal and Qi (2017) found that XGBoost outperformed several traditional machine learning algorithms in stock market prediction tasks. Similarly, Fischer and Krauss (2018) reported that boosting-based models achieved strong forecasting performance when applied to financial time series data. Despite its advantages, XGBoost requires careful hyperparameter tuning and may become computationally intensive when dealing with large datasets (Chen and Guestrin, 2016).

Nevertheless, its high predictive accuracy and robustness have made it one of the most widely adopted machine learning algorithms in financial forecasting research.

2.1.6 Gradient Boosting Machine Learning model

Gradient Boosting is an ensemble machine learning technique that combines multiple weak learners, typically decision trees, to create a strong predictive model. Proposed by Friedman

(2001), the algorithm builds trees sequentially, with each new tree focusing on correcting the prediction errors of the previous ensemble. Unlike Random Forest, which constructs trees independently, Gradient Boosting learns iteratively, allowing it to capture complex patterns and non-linear relationships within data. Due to these characteristics, Gradient Boosting has been widely applied in financial forecasting and predictive analytics. Several studies have demonstrated the effectiveness of Gradient Boosting in financial markets. For example, Natekin and Knoll (2013) highlighted that Gradient Boosting often achieves superior predictive accuracy compared with traditional machine learning methods because of its ability to minimize prediction errors through iterative learning. In a financial context, Jiang (2021) found that boosting-based algorithms effectively captured market trends and improved stock price prediction performance.

However, despite its strong predictive capability, Gradient Boosting may be sensitive to hyper parameter settings and can suffer from over fitting if the model becomes excessively complex (Friedman, 2001). Nevertheless, its ability to model intricate relationships within financial data has made it one of the most widely used ensemble learning techniques in forecasting applications.

2.2 The Foundations of Technical Indicators in Financial Forecasting

Technical analysis is a widely known approach in financial markets, which applies historical charts and statistical data to forecast future price movements in financial markets.

Market psychology and investor behavior tend to repeat over time, which is the main assumption of technical analysis. Based on that, it identifies trends and price patterns (Sezer, Gudelek and Ozbayoglu 2020). In their simple technical trading rules paper, professors Brock, Lakonishok and LeBaron (1992) provided foundational empirical evidence which supports technical trading rules. They claim that statistically significant returns can be generated by simple moving average strategies. Moving Averages (MA), Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD) and Bollinger Bands are the most widely known technical indicators in the financial market. Moving averages are basic technical indicators that smooth out price data using historical price data, which helps identify trend direction and detect potential support and resistance levels. On the other hand, Relative Strength Index (RSI) is a momentum indicator, which evaluates the speed and magnitude of price changes (Fernando, 2025). Besides, MACD gives signals for potential reversals by evaluating the relationship between two moving averages of an asset’s price. Furthermore, Bollinger Bands measures market volatility using a formulaic method and helps detect overbought or oversold conditions. It uses a simple moving average and standard deviations. According to Sezer, Gudelek and Ozbayoglu (2020), these indicators are still crucial in modern financial trading. They claim that those indicators play a vital role in financial time series analysis. This is in consistent with findings of research by Amini and Kalantari (2024) that technical indicators, RSI and MACD in particular, improve forecasting accuracy when used in combination with machine learning frameworks.

Furthermore, Chong and Ng (2008) found that the effectiveness of technical indicators is highly sensitive to market conditions. Likewise, Neely et al. (2014) also did similar research and found that technical indicators can give more accurate forecasting results when used with other analytical modules. This finding is also consistent with that of a more recent study conducted by

Sezer et al and Gu, et al (2020) that rather than using standalone, technical indicators provide better performance results when used as feature engineering tools within data-driven models.

Therefore, technical indicators play a key role in evaluating crucial aspects of market behavior such as trend, momentum and volatility, which develops a basis for hybrid forecasting models combining traditional analysis with machine learning models.

2.3 Limitations of Traditional Technical Indicators

Even though technical indicators are widely used in financial analysis, their effectiveness and accuracy are very sensitive to changes in financial market, especially when used standalone.

Technical indicators use historical price data and may fail to respond effectively to unexpected market changes, especially in highly volatile markets, thus, giving false signals (Gupta, 2025). Similarly, the research conducted by Chong and Ng (2008) found that when the market is highly volatile, technical indicators generate false and inaccurate signals. Take RSI as an example. This is a momentum-based indicator whose practical reliability decreases when it detects overbought or oversold conditions without a corresponding price reversal. They also claim that accuracy of those indicators is highly sensitive to changes in market conditions and the selection of parameters. This is in harmony with findings of research by Sezer, Gudelek and Ozbayoglu (2020) that technical indicators often suffer from instability and provide less accurate forecasting results during periods of high market volatility. Furthermore, Lo, Mamaysky and Wang (2000) claim that technical indicators often fail to capture complex, non-linear relationships in financial markets. In other words, they mainly use mathematical formulas, hindering them from adapting to changing market conditions. They also claim that simple rule-based strategies often fail to capture complicated, non-linear relationships and noise in gold price structure. Furthermore, according to the research carried out by Gu, Kelly and Xiu (2020), machine learning techniques can handle high-dimensional and non-linear relationships better than rule-based traditional methods. They also found that technical indicators used in isolation often give conflicting market signals. In other words, various technical indicators might generate opposing or conflicting signals simultaneously, making it hard for financial traders to make appropriate trading decisions. To resolve this issue, Neely et all (2014) offers a more integrated and systematic approach to effectively use technical indicators. They suggest using technical indicators as input variables within more advanced forecasting models rather than using them in isolation, providing more accurate forecasting results. This is also in consistent with the findings of research by Sezer et al., (2020); Konur, Göçken and Dosdoğru, (2024) that combining technical indicators with more complicated machine learning models significantly improves forecasting performance (Sezer et al., 2020; Konur, Göçken and Dosdoğru, 2024).

2.4 The Evolution of Machine Learning Models: Random Forest, Gradient Boosting and XGBoost

The growing complexity of financial markets has encouraged researchers to adopt machine learning techniques capable of identifying non-linear relationships and hidden patterns within large financial datasets. Traditional statistical models often rely on restrictive assumptions regarding data distributions and linearity, whereas machine learning algorithms can learn complex interactions directly from the data. Among the most influential developments in this area are ensemble learning methods, particularly Random Forest, Gradient Boosting, and XGBoost.

Random Forest, introduced by Breiman (2001), represented a major advancement in predictive modelling through the aggregation of multiple decision trees constructed using bootstrap sampling and random feature selection. By combining the predictions of numerous trees, Random Forest reduces variance and improves generalisation performance, making it particularly suitable for noisy financial datasets. Several studies have demonstrated its effectiveness in financial forecasting. For example, Patel et al. (2015) reported that Random Forest outperformed several traditional machine learning techniques in predicting stock market movements, while Gu, Kelly and Xiu (2020) highlighted the algorithm's ability to capture complex relationships within financial data. Building upon ensemble learning principles, Gradient Boosting was proposed by Friedman (2001) as a sequential learning framework in which each new decision tree is trained to correct the prediction errors of previously constructed trees. This iterative optimisation process allows Gradient Boosting to model highly complex and non-linear patterns. Due to its strong predictive capability, Gradient Boosting has been widely applied in forecasting applications where accuracy is prioritised over model simplicity. Natekin and Knoll (2013) argue that boosting algorithms frequently achieve superior predictive performance because they continuously minimise residual errors throughout the learning process. A further advancement emerged with the development of Extreme Gradient Boosting (XGBoost) by Chen and Guestrin (2016). XGBoost extends the Gradient Boosting framework through regularisation mechanisms, parallel processing, and computational optimisations designed to improve both predictive performance and scalability.

These enhancements have made XGBoost one of the most widely adopted algorithms in machine learning competitions and financial forecasting applications. Studies such as Fischer and Krauss (2018) have demonstrated the effectiveness of boosting-based algorithms in modelling financial market behaviour, while Sezer, Gudelek and Ozbayoglu (2020) identified ensemble learning approaches as among the most promising techniques for financial time-series prediction.

The progression from Random Forest to Gradient Boosting and subsequently XGBoost reflects the broader evolution of machine learning towards increasingly sophisticated ensemble architectures capable of handling the complexity, volatility, and non-linearity inherent in financial markets. Consequently, these algorithms have become central to modern forecasting research and provide a strong foundation for developing hybrid forecasting frameworks.

2.5 Hybrid Forecasting Frameworks: A New Synthesis

The limitations of standalone forecasting methods have led researchers to develop hybrid forecasting frameworks that integrate multiple analytical techniques within a single predictive system. Hybrid forecasting is based on the premise that no single model can fully capture all dimensions of financial market behaviour. Instead, combining complementary methods enables researchers to exploit the strengths of different approaches while mitigating their individual weaknesses. In financial forecasting, hybrid frameworks commonly integrate technical indicators, statistical methods, and machine learning algorithms. Technical indicators such as RSI, MACD, and Bollinger Bands provide information regarding market momentum, trend direction, and volatility conditions, while machine learning models are capable of identifying complex non-linear interactions among these variables. By combining these approaches, hybrid models can generate more robust predictions than either component could achieve independently. Empirical evidence supports the effectiveness of hybrid forecasting approaches. Kim (2003) demonstrated that combining technical analysis with machine learning techniques improved stock market prediction accuracy compared with conventional methods. Similarly, Patel et al. (2015) found that hybrid systems integrating technical indicators and machine learning algorithms consistently outperformed standalone forecasting models. More recently, Sezer, Gudelek and Ozbayoglu (2020), in their comprehensive review of financial forecasting research, concluded that hybrid and ensemble architectures generally achieve superior performance because they can capture diverse market signals simultaneously. Furthermore, hybrid frameworks are particularly valuable in highly volatile markets such as gold trading. Gold prices are influenced by momentum effects, volatility fluctuations, investor sentiment, macroeconomic conditions, and geopolitical uncertainty. As a result, relying on a single forecasting methodology may overlook important market information.

Hybrid models address this challenge by synthesising multiple sources of predictive information into a unified framework. According to Krollner, Vanstone and Finnie (2010), combining multiple forecasting approaches often enhances model robustness and improves predictive performance under changing market conditions. Consequently, hybrid forecasting frameworks have emerged as an important research direction in financial prediction and provide a promising foundation for improving gold price forecasting accuracy.

2.6 Gaps in the Literature

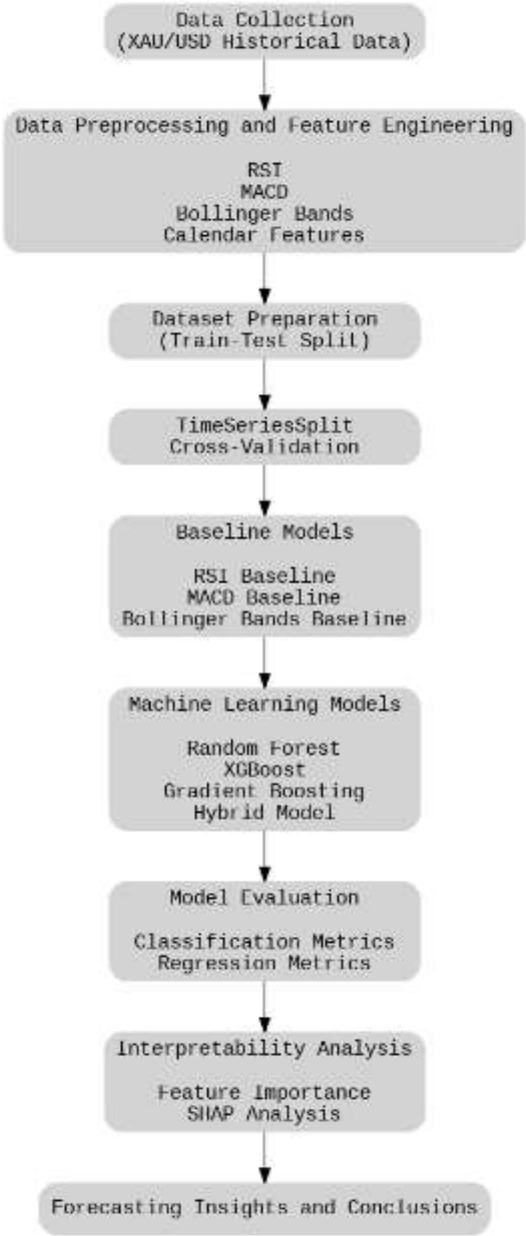
Although previous studies have extensively examined the use of technical indicators and machine learning algorithms in financial forecasting, several important research gaps remain. First, many studies focus primarily on equity markets, cryptocurrencies, or foreign exchange markets, while comparatively limited attention has been given to gold price forecasting using hybrid ensemble learning approaches. Second, existing research frequently evaluates machine learning models such as Random Forest, XGBoost, or Gradient Boosting individually, with relatively few studies investigating how combinations of these algorithms can improve forecasting performance. Third, while technical indicators including RSI, MACD, and Bollinger Bands are widely used in trading practice, limited research has systematically compared their standalone forecasting performance against advanced machine learning and hybrid forecasting frameworks within a single study. Finally, many previous studies prioritise predictive accuracy without providing sufficient model interpretability, making it difficult to understand the contribution of individual indicators to forecasting outcomes. To address these gaps, this study develops a hybrid forecasting framework that integrates technical indicators with ensemble machine learning algorithms for gold price prediction. The study not only compares traditional technical indicator models with Random Forest, XGBoost, Gradient Boosting, and hybrid models, but also employs Feature Importance and SHAP analysis to improve model interpretability and provide deeper insights into the factors influencing gold price movements.

3. Methodology

3.1 Research Design

This is a quantitative, experimental research conducted to measure the forecasting performance of a hybrid machine learning framework for XAU/USD price movements. All data analysis in this study is done using Google Colab, which is a free, cloud-based Jupyter Notebook environment and several analytical libraries, such us Pandas, NumPy, Sklearn, XGBoost, Matplotlib, SHAP and others were applied to do the analysis.

Diagram 4 Methodological framework integrating technical indicators with machine learning models for gold price prediction



According to the diagram 4, the research framework is structured progressively and has several analytical stages.

It starts with gold price data collection, followed by data preprocessing and feature engineering where the chosen data was cleaned and the selected technical indicators, such as RSI, MACD and Bollinger Bands in combination with calendar features were calculated.

The next step was dataset preparation, in which stage, the processed data was subject to Train-Test split. This was followed by TimeSeriesSplit stage where the 5-fold cross validation was conducted with Random Forest, XGBoost and Gradient Boosting machine learning algorithms. In the next stage, RSI, MACD, and Bollinger Bands baseline models were developed as one of the main objectives of the study is to develop baseline models using these individual technical indicators to establish a performance benchmark. After that, machine learning models, such as Random Forest, XGBoost, Gradient Boosting and Hybrid model were implemented, followed by model evaluation stage, where the classification metrics and regression metrics were computed.

Also, Feature Importance and SHAP analysis were conducted to find out how each individual input variable contributed in the implementation of the machine learning algorithms. Finally, the research provides forecasting insights and conclusions based on data analysis results. Now, we will discuss each stage in very detail.

3.2 Data collection

This study uses secondary data consisting of the historical price data of XAU/USD (Gold/US Dollar) pair, collected from Myfxbook API , which is a reliable financial source, widely used by financial traders. The data is sourced from the following website: https://www.myfxbook.com/forex-market/currencies/XAUUSD-historical-data. Moreover, the timeframe of the dataset is daily data because they provide sufficient granularity for technical analysis and they have relatively stable market structure in comparison to higher frequency data frames.

Diagram 5. Snapshot of daily price data of XAU/USD (Gold/US Dollar) pair utilized in this study.

	Date	Open	High	Low	Close	Change(Pips)	Change(%)
0	3/1/2026 0:00	5351.51	5392.49	5332.91	5387.03	3552	0.66
1	2/27/2026 0:00	5188.23	5280.82	5166.80	5279.95	9172	1.74
2	2/26/2026 0:00	5165.12	5205.51	5130.30	5184.90	1978	0.38
3	2/25/2026 0:00	5141.52	5217.73	5120.31	5164.94	2342	0.45
4	2/24/2026 0:00	5230.95	5249.73	5093.33	5143.71	-8724	-1.70

We tried to collect at least 5 year daily price data of gold, but we could find only almost 3 year daily price data, extending from July 25, 2023, to March 1, 2026, which is the only existing historical daily price of gold in the market. Table 1 shows that the dataset has 7 variables, including Date, Open price, High price, Low price, Closing price, Change (Pips), and Change (%).

However, in this research, only two variables, such as Date and Closing price were utilized.

3.3 Data preprocessing and feature engineering

The dataset went through data preprocessing stages using Power Query to enhance the quality of the dataset to ensure that the raw data is structurally compatible with the Random Forest, XGBoost and Gradient Boosting machine learning algorithms. This stage is vital because incomplete financial data can have a significant negative impact on forecasting performance.

Thus, the collected data was cleaned by detecting and removing missing values, duplicated observations and inconsistencies where necessary. In the collected dataset, there were non-tabular titles and blank columns, so the first cleaning process involved removing irrelevant top rows as they would influence data analysis accuracy later. Also the first column, Date, was decomposed by removing redundant time component (00:00). Besides that, as you can see in the Diagram 6, the

Date variable was separated into three distinct integer-based variables, such as Year, Month and Day as it allows us to find out whether cyclical or seasonal effects influence gold price behavior.

Diagram 6. Decomposition of Date variable and Technical indicators calculation

se	Change(Pips)	Change(%)	Year	Month	Day	RSI	MACD	MACD_Signal	BB_High	BB_Low
32	7064	1.46	2026	1	21	42.526874	-39.946594	-40.529254	5351.199364	4631.875636
55	8845	1.86	2026	1	20	40.555416	-55.554403	-43.534284	5355.820140	4611.429860
50	-703	-0.15	2026	1	19	37.823423	-74.971297	-49.821687	5356.357878	4569.000122
54	3861	0.83	2026	1	18	38.145612	-88.767995	-57.610948	5357.105448	4532.985552
65	-1452	-0.32	2026	1	16	35.954240	-104.699447	-67.028648	5357.244901	4486.592099

After data preprocessing procedures, the study underwent feature engineering procedures to create explanatory variables from historical gold price movements. The following technical indicators were chosen for the study, including Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), MACD Signal Line, Bollinger Bands Upper Band (BB_High), and Bollinger Bands Lower Band (BB_Low). Let us define each technical indicator one by one. So, the first chosen indicator is the Relative Strength Index (RSI), which can take values from 0 to 100, and is used to capture market momentum and identify overbought and oversold market conditions. According to Fidelity Investments (2026), traditionally, the value of RSI above 70 means that an asset is considered overbought, giving a signal of price downward trend, and when it is below 30, it is considered as oversold, giving a signal of price upward trend.

Moreover, other technical indicator used in this model are the Moving Average Convergence Divergence (MACD) and MACD Signal Line. The former is mainly used to measure the strength of trend and potential trend reversals, detecting bullish and bearish momentum transitions, while the latter is used to smooth MACD fluctuations and help in finding crossover patterns. Fidelity Investments (2026) claim that the MACD value exceeding 0 is considered bullish whereas the value crossing below 0 is considered bearish. Furthermore, Bollinger Bands technical indicator was also used in this model. It is used to capture market volatility conditions.

Bollinger Bands have upper (BB_High) and lower (BB_Low) bands, which are calculated above and below a simple moving average using standard deviation based volatility ranges. This indicator is used to find out whether prices are high or low on a relative basis. Fidelity Investments (2026) claim that when the price reaches upper or lower bands, there is a likelihood of a sharp price movement in either direction. They also claim that when the price moves out of the band, there is a strong trend continuation. So, these momentum, trend following and volatility measures are computed to be used as explanatory variables in our hybrid model with the combination of machine learning models, capable of capturing different aspects of XAU/USD market dynamics.

3.4 Dataset Preparation Train-Test split

The next step is the dataset preparation, in which stage, we transformed the dataset into a supervised machine learning structure. The technical indicators, RSI, MACD and Bollinger Bands and calendar features were used as explanatory variables in this study, while the primary target variable is the directional movement of the price, defined as a binary variable. To be more specific, the value of 1 is given for price up movement, and the value of 0 is given for price down movement. This task evaluates the practical utility of the model for trend-following strategies, where the sign of the move is often more critical than its exact magnitude. Using the 80/20 train-test split method, we separated the dataset as training and testing subsets. In other words, in the dataset, we have 572 training data rows, used for model learning, and 143 testing data rows, used for out-of-sample performance evaluation. This categorization was important to measure the generalization capability of the machine learning algorithms on future market observations.

3.5 TimeSeriesSplit and 5-Fold Time series cross validation

In order to ensure the reliability and generalizability of forecasting results, the study conducted a 5-fold time series cross validation method, minimizing the risk of overfitting.

TimeSeriesSplit keeps chronological ordering within financial data, hence providing a more realistic simulation of real-world forecasting conditions, making it different from traditional random cross-validation methods. The utilization of TimeSeriesSplit is vital because financial time-series data have temporal dependence, indicating that future price movements are affected by historical market behavior. In order to do this 5-fold cross validation, the historical dataset is divided into five discrete subsets. So, how the model works is that the models are iteratively trained on four of these subsets while the fifth is reserved as an independent test set. This is repeated five times in a row, making sure that each data point is applied for both training and validation. By averaging the results across these five iterations, the study reduces the risk of overfitting, where a model becomes too specialized to a specific segment of data and fails to perform on new, unseen market conditions. So, the 5-fold time series cross validation is conducted with machine learning algorithms, including Random Forest, XGBoost and Gradient Boosting.

3.6 Baseline Models

In earlier feature engineering stage, technical indicators, including Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands, were calculated later to be used as input variables for machine learning models. Now, at this baseline models stage, these indicators are computed to see how accurate each individual technical indicator forecasts the gold price as standalone baseline models, serving as performance benchmarks for machine learning models.

So, these baseline models follow certain rules and mathematical formulas to predict the direction of gold prices. Let us discuss how each baseline model is constructed. As it was discussed in the previous sections, the Relative Strength Index (RSI) shows the momentum of asset prices, detecting overbought and oversold conditions. So, this model triggers a buy signal when its value goes down below 30, showing an oversold condition. On the other hand, it gives a sell signal when its value exceeds 70, showing an overbought condition. The second technical indicator is the trend-detecting MACD model, which finds the difference between short term and long term moving averages. So, when the difference is positive, it triggers a buy signal, and when it is a negative, it triggers a sell signal. Furthermore, Bollinger Bands indicator shows the volatility of asset prices. This baseline gives a buy or sell signal when the price crosses either upper or lower bands. In other words, when the price reaches lower band level, it is considered oversold condition, triggering a buy signal. On the contrary, when the price reaches the upper band level, it is considered overbought state, triggering a sell signal. So, this is the logic behind how each individual baseline model triggers a buy or sell signal. The performance of each baseline model was measured by their overall accuracy level, precision, recall, f1-score and support outcomes.

3.7 Machine Learning Models

In this stage, the role of technical indicators shifts from triggering buy or sell signal to serving as input variables in the development of machine learning algorithms. In this study, three widely known ensemble architectures, namely Random Forest, XGBoost and Gradient Boosting, are utilized to process technical time-series data. These algorithms are applied to detect complex, non-linear relations and volatility patterns that traditional baseline models fail to capture. These ensemble models are selected based on existing literature and for their proven efficacy with tabular financial data and their ability to handle high-dimensional feature sets without the prohibitive data requirements of deep neural networks. First, each of those ensemble models is computed individually and tested for their predicting accuracy. Each model's performance was evaluated by their overall accuracy level, precision, recall, f1-score and support outcomes. After doing individual module evaluation, a hybrid forecasting model was developed based on the classification performance results of each individual machine learning model. So, first, the classification results of XGBoost and Gradient Boosting were compared. Then based on the results, either XGBoost and Gradient Boosting or both were applied to boost random forest machine learning results, which combines the ensemble properties of Random Forest, XGBoost and Gradient Boosting algorithms to perform automated feature weighting. Fischer and Krauss (2018) and Neely et al. claim that the combination of different technical signals through machine learning can provide a better forecasting outcome compared to traditional methods. The performance of the final hybrid framework is benchmarked against the performance of the constructed baseline models, thus providing empirical validation regarding the comparative efficacy of algorithmic forecasting versus traditional rule-based trading methods in the gold market.

3.8 Model Evaluation Metrics

The study applied classification and regression metrics to evaluate the forecasting performance of both baseline models and machine learning models. Regarding the classification metrics, overall Accuracy Level, Precision, Recall and F1-Score were used to compare the predicting performance of the traditional models with that of machine learning models. Overall Accuracy Level evaluated the proportion of accurately classified observations compared to the total number of predictions, while the Precision measured the reliability of positive predictions. Moreover, Recall evaluated how accurately each model generated positive price movements, whereas the F1-Score provided a balanced measure between Precision and Recall. On the other hand, regression based predicting performance of the models were evaluated through Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2).

RMSE shows the square root of average squared forecasting errors, while MAE shows the average magnitude of prediction mistakes in absolute terms. Finally, R^2 measured the significance of the models in predicting gold prices in this model.

3.9 Feature Importance and SHAP Analysis

Feature Importance analysis was conducted individually for Random Forest, XGBoost and Gradient Boosting ensemble models to identify the relative contribution of each explanatory variable to their forecasting performance. The Mean Decrease in Impurity (MDI) scores given in the Feature Importance chart (as you can see in results section) shows the ranking of technical indicators based on their contribution on model's predictive performance. Furthermore, SHAP analysis was carried out to explain how individual variables affected model outputs across observations. The primary advantage of SHAP analysis over Feature Importance analysis is that SHAP values provide directional interpretation by indicating whether variables positively or negatively contribute to predictions. These analyses are vital because the study not only measured predicting performance of the models, but also investigated the economic and financial meaning behind model decisions.

4. Results and Discussion

4.1 Five-Fold Cross Validation Results

Diagram 7. Five-Fold Time-Series Cross Validation

```
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:52.94%
XGBoost Mean CV Accuracy50.42%
Gradient Boosting Mean CV Accuracy53.78%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:45.80%
XGBoost Mean CV Accuracy43.70%
Gradient Boosting Mean CV Accuracy48.32%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:49.58%
XGBoost Mean CV Accuracy46.78%
Gradient Boosting Mean CV Accuracy49.30%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:48.53%
XGBoost Mean CV Accuracy46.01%
Gradient Boosting Mean CV Accuracy50.84%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:50.59%
XGBoost Mean CV Accuracy47.73%
Gradient Boosting Mean CV Accuracy51.43%
================================================================
```

The diagram 7 shows the results of five-fold-time-series cross-validation for the Random Forest, XGBoost, and Gradient Boosting models. According to the results, across the five validation folds, the models achieved mean accuracies ranging between approximately 43% and 54%. Amongst these machine learning models, the Gradient Boosting consistently demonstrated the strongest performance, achieving the highest accuracy in four out of the five folds. Its accuracy varied from 48.32% to 53.78%, indicating relatively stable predictive performance across different time periods. In contrast, the Random Forest produced similar results, with accuracies ranging from 45.80% to 52.94%, while XGBoost recorded the lowest performance, with accuracies between 43.70% and 50.42%. in order to compare the overall performance, the average accuracy across all folds was calculated as you can see in the diagram. The results show that the Gradient Boosting provides the most reliable predictive capability among the three algorithms. However, the relatively modest accuracy levels suggest that predicting the target variable remains challenging. The decline in performance from the first fold to subsequent folds may indicate that the underlying data-generating process changes over time, reducing the ability of the models to generalize to future observations. Such behavior is common in financial and economic datasets, where market conditions, borrower behavior, or macroeconomic factors evolve over time.

Furthermore, the variation in accuracy across folds demonstrates that model performance is sensitive to the specific time period used for validation, highlighting the importance of employing time-series cross-validation rather than a single train-test split, as it provides a more comprehensive evaluation of model stability and robustness.

4.2. Classification Performance Results of Traditional Baseline Models and Discussions

Classification metrics were conducted for baseline models as discussed in the methodology section. Now, let us discuss the results of each baseline model classification performance in very detail.

Diagram 8 Relative Strength Index (RSI) Baseline Model Classification Report

```
=======================RSI BASELINE MODEL====================
Accuracy: 50.41%

Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.03      0.06        63
           1       0.50      1.00      0.66        60

    accuracy                           0.50       123
   macro avg       0.75      0.52      0.36       123
weighted avg       0.75      0.50      0.35       123

=============================================================
```

According to the diagram 8, overall accuracy level of RSI baseline model is 50.41%, which indicates that the model accurately forecasted the next-day gold price direction in almost half of the observations. Furthermore, the classification report shows that for class 0 (Price Decrease), the precision value is 1.00, meaning that whenever the model forecasted a price decrease, it was always correct, but the recall value is just 0.03, indicating that the model identified only 3% of all actual price decreases. In other words, the RSI model almost never forecasted price decreases. Moreover, the F1-Score is also quite low, only 0.06, indicating poor performance in detecting bearish market states. Even though the model predicted a few downward price movements correctly, it failed to capture the vast majority of actual price decreases in gold prices. On the other hand, in the case of class 1 (Price Increase), precision value of 0.5 means that the RSI model predicted price increases correctly only in half of the cases even though the recall value of 1.00 shows that the model successfully identified all actual price increases.

This implies that the RSI model tend to classify almost every price change as a price increase, therefore, labeling many price decreases as price increases. However, the F1-Score of the model is 0.66, significantly higher than that of Class 0, which indicates that the RSI model forecasts price increases more accurately than price decreases. Furthermore, the macro average and weighted average metrics show overall classification performance across both classes. The F1-Score of 0.36 means that the RSI model struggles to distinguish between bullish and bearish market conditions, meaning that the RSI model has limited forecasting capability as a standalone forecasting model.

Diagram 9. MACD Baseline Model Classification Report

```
====================MACD BASELINE MODEL======================
Accuracy: 44.06%

Classification Report:
              precision    recall  f1-score   support

           0       0.51      0.41      0.45       404
           1       0.39      0.48      0.43       311

    accuracy                           0.44       715
   macro avg       0.45      0.45      0.44       715
weighted avg       0.45      0.44      0.44       715

============================================================
```

The diagram 9 shows that the MACD baseline model obtained an overall classification accuracy of 44.06%, indicating that only 44.06% of the cases, the model correctly forecasted the direction of gold prices. Besides that, regarding the analysis of Class 0 (Price Decrease), it achieved a precision value of 0.51, meaning that almost half of the price decrease predictions of the model were correct, but the recall value of 0.41 shows that the model could capture only 41% of the actual price decreases. Also, the F1-Score of 0.45 means that the MACD suffers to detect a substantial proportion of actual price downward movements. Regarding the analysis of Class 1(Price Increase), the precision value of 0.39 shows that only 39% of the predicted upward price movements were correct, while the recall value of 0.48 indicates that the model could successfully capture 48% of actual price increases. The F-1 Score of 0.43 means that the MACD model suffers to predict bullish market conditions. Almost half of the actual price upward movements were missed and also more than half of the forecasted price increases were incorrect. Furthermore, the macro average F-1 score of 0.44 indicates that the MACD model suffers to distinguish effectively between increasing and decreasing gold prices. The MACD, as a standalone model, performs weakly in forecasting gold prices.

Diagram 10. Bollinger Bands Baseline Model Classification Report

```
=====================BOLLINGER BANDS BASELINE=====================
Accuracy:57.53%

Classification Report:
              precision    recall  f1-score   support

           0       0.78      0.34      0.47        41
           1       0.51      0.88      0.64        32

    accuracy                           0.58        73
   macro avg       0.64      0.61      0.56        73
weighted avg       0.66      0.58      0.55        73

==================================================================
```

According to the diagram 10, the Bollinger Bands achieved an overall accuracy score of 57.53%, meaning that the model accurately predicted the direction of price movements of gold prices in almost 58 out of 100 cases, which is quite a good result. Regarding the analysis of Class 0 (Price Decrease), the precision value of 0.78 indicates that almost 78% of the forecasted downward movements were correct, but the recall score of 0.34 shows that the model could successfully recognized only 34% of actual price decreases. This means that although the model accurately predicts price decreases, it fails to capture many actual bearish movements. The F1-score of 0.47 means that the model misses actual price decreases in almost half of the cases. From this, it is implied that the Bollinger Bands prefers to avoid giving false alarms, resulting in high precision but low recall values. Furthermore, in the case of the Class 1 (Price Increase), the precision value of 0.51 indicates that approximately half of the all forecasted price increases were correct, while the recall value of 0.88 shows that the model could successfully capture almost 88% of actual bullish market movements. Besides, the F-1 score of class 1 is 0.64, considerably higher than that of Class 0, meaning that the model performs better in predicting bullish market movements than bearish ones. Macro average and weighted average classification outcomes show that the model has a more balanced classification performance, meaning that in almost more than half of the cases, the model can capture upward and downward price movements successfully.

4.3. Classification Performance Results of Machine Learning Models and Discussions

Diagram 11. Random Forest Classification Results

```
================ RANDOM FOREST RESULTS =====================
Overall Test Accuracy:  59.44%

Detailed Classification Report:
              precision    recall  f1-score   support

           0       0.61      0.92      0.73        86
           1       0.46      0.11      0.17        57

    accuracy                           0.59       143
   macro avg       0.53      0.51      0.45       143
weighted avg       0.55      0.59      0.51       143

============================================================
```

The Diagram 11 shows that the Random Forest ensemble method obtained an overall classification accuracy of 59.47%, meaning that in almost 60% of the cases, the model predicts gold price direction correctly. Furthermore, for Class 0(Price decrease), the precision value of 0.61 means that almost 61% of the predicted downward price movements are correct, while the recall value of 0.92 indicates that the model can capture almost 92% of actual price decreases. Also, the F1-score of 0.73 clearly shows that the model can successfully capture bearish market movements in most of the cases. However, in the case of Class 1(price increase), the precision value of 0.46 means that more than half of the predicted price increases are incorrect. The recall value of 0.11 indicates that the model misses almost 90% of the actual price increases. The F1-score of 0.17 shows that the Random Forest Machine learning model suffers to capture bullish market movements. From this, it is implied that the model is better at predicting downward price movements than upward price movements. Moreover, the weighted average f1-score of 0.51 means that the model has moderate level of forecasting gold price direction.

Diagram 12. XGBoost Classification Results

```
=====================XGBOOST RESULTS ========================
Overall Test Accuracy:  54.55%

Deatiled Classification Report:
              precision    recall  f1-score   support

           0       0.62      0.65      0.63        86
           1       0.42      0.39      0.40        57

    accuracy                           0.55       143
   macro avg       0.52      0.52      0.52       143
weighted avg       0.54      0.55      0.54       143

============================================================
```

According to the diagram 12, the overall classification accuracy of XGBoost machine learning is 54.55%, indicating that almost 55% of the model's gold price direction predictions are correct. For Class 0 (price decrease) analysis, the precision value of 0.62 means that when the XGBoost predicts a downward price direction, almost 62% of them are correct, while the recall of 0.65 means that it could successfully capture around 65% of the actual price decreases. The f1-score of 0.63 means that the model is proficient at recognizing bearish market movements.

Regarding the Class 1(price increase), the precision value of 0.42 means that only around 42% of the forecasted price increases are correct and the recall value of 0.39 means that the model misses price increases in almost 60 out of 100 cases. This is justified by the f1-score of 0.4, showing that the model has higher tendency in recognizing bearish market movements than bullish market movements. The weighted average f1-score shows that the model has moderate level of gold price forecasting capability when used in isolation.

Diagram 13. Gradient Boosting Classification Results

Gradient Boosting Results
Accuracy: 55.94%
Classification Report:

	precision	recall	f1-score	support
0	0.62	0.69	0.65	86
1	0.44	0.37	0.40	57
accuracy			0.56	143
macro avg	0.53	0.53	0.53	143
weighted avg	0.55	0.56	0.55	143

According to the diagram 13, the overall classification accuracy of Gradient Boosting model is 55.94%, meaning that the model is overall slightly outperforming XGBoost model in predicting gold price directions, which might be due the fact that Gradient Boosting performs better than XGBoost when the sample size of the data is small. So, regarding the Class 0 (price decreases), the precision value of 0.62 is equal to that of XGBoost, meaning that in around 62% of the cases, both models correctly predict downward price movements.

However, the recall value of Gradient Boosting is slightly higher, 0.69, which shows that the model can capture almost 70% of the actual price decreases, thus leading to a higher f1-score of 0.65. On the contrary, in the case of class 1(price increase), the precision value of Gradient Boosting model is 0.44, slightly higher than that of XGBoost, but it is quite a low result, meaning that in almost 56 percent of the cases, the Gradient Boosting model makes wrong price increase predictions. The recall of 0.37 means that the model misses around 63% of the actual price increases. The f1-socre value of 0.4 shows that the model suffers to distinguish between gold price upward and downward movements. However, the weighted average f1-score of 0.55 shows the balanced classification performance of the model.

Diagram 14. Hybrid Model 1 (RF, XGB, and GB) Classification Results

Hybrid Model Results (Random Forest, XGBoost and Gradient Boosting)
Accuracy: 59.44%
Classification Report:

	precision	recall	f1-score	support
0	0.61	0.88	0.72	86
1	0.47	0.16	0.24	57
accuracy			0.59	143
macro avg	0.54	0.52	0.48	143
weighted avg	0.56	0.59	0.53	143

The diagram 14 shows the classification results of the hybrid model developed by integrating features of those three machine learning models, including Random Forest, XGBoost and Gradient Boosting. So, we can see that the overall classification accuracy of hybrid model is equal to that of Random Forest model when used in isolation, which means that the integration of both XGBoost and Gradient Boosting models into the hybrid model did not boost the overall classification accuracy level. However, the precision value for Class0 (price decrease) did not change, but the recall value slightly decreased from 0.92 (Random Forest) to 0.88 (Hybrid model).

On the other hand, the precision, recall and f1-socre values for Class 1(price increase) has slightly increased in hybrid model. It is implied that the hybrid model is shifting the focus from identifying bearish market movements to detecting more bullish market movements, but the effect is quite little. As a result, the weighted average and macro average f1-score values of hybrid module increased slightly compared to random forest outcomes. Now, we will test how excluding Gradient Boosting model from the hybrid model can affect the forecasting framework of the hybrid model.

Diagram 15. Hybrid model 2 (Random Forest and XGBoost) Classification Results.

```
Hybrid Model Results (Random Forest and XGBoost)
Accuracy: 57.34%
Classification Report:
              precision    recall  f1-score   support

           0       0.61      0.80      0.69        86
           1       0.43      0.23      0.30        57

    accuracy                           0.57       143
   macro avg       0.52      0.52      0.50       143
weighted avg       0.54      0.57      0.54       143
```

The diagram 15 shows that exclusion of the Gradient Boosting model from the hybrid model had a negative impact on the overall classification accuracy of the hybrid model, 57.34%.

Moreover, the recall value for Class 0 (price decrease) slightly decreased to 0.80, which is still quite a good result. Talking about the positive impacts, so the model has increased the recall and f-1 scores for Class 1(price increases), but it decreased the precision value for class 1, meaning that after removing the Gradient Boosting model from the hybrid model, the new hybrid model made more errors in forecasting gold price increases although it started to capture more of the actual price increases. However, the macro and weighted averages show better results for the new hybrid model.

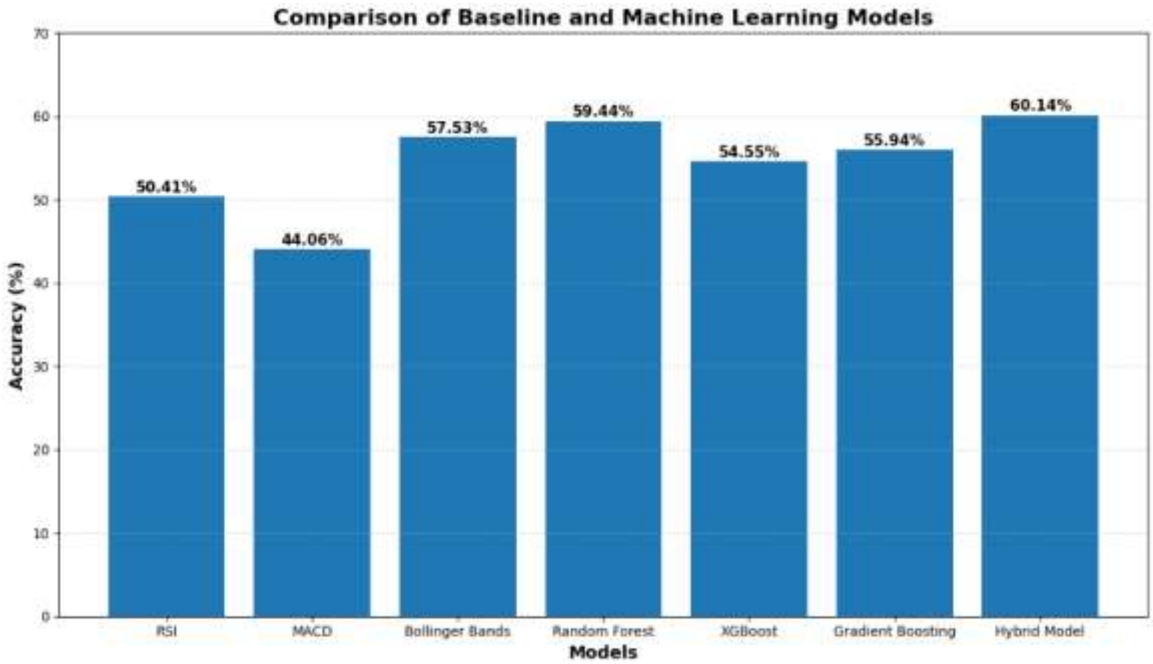
As it was discussed earlier that the Gradient Boosting model performed better than XGBoost mode while used independently in forecasting gold price predictions. So, now we will create another hybrid model by integrating the properties of only Random Forest and Gradient Boosting models. Diagram 16. Hybrid model 3(Random Forest and Gradient Boosting) Classification Results

```
Hybrid Model Results (Random Forest and Gradient Boosting)
Accuracy: 60.14%
Classification Report:
              precision    recall  f1-score   support

           0       0.61      0.93      0.74        86
           1       0.50      0.11      0.17        57

    accuracy                           0.60       143
   macro avg       0.56      0.52      0.46       143
weighted avg       0.57      0.60      0.51       143
```

The Diagram 16 illustrates that the overall classification accuracy of the hybrid model (60.14%) has noticeably improved after removing XGBoost from the model, meaning that the hybrid model's ability to predict gold prices has increased considerably. One of the possible reasons might be that Gradient Boosting performs better with small size data compared to XGBoost model. So, this overall classification.

Moreover, the recall and f1-socre values for class 0 (price decrease) increased, indicating that the hybrid model can now recognize more of the actual downward market movements even though those for class 1 (price increase) did not change. Also, this exclusion of XGBoost from the model increased the precision for Class 1, meaning that the model is now providing more accurate forecasting results for price increases. Furthermore, the macro and weighted average f1-score values show that the hybrid model now has much better balance in predicting overall gold price direction.

Diagram 17. Comparing Baseline and Machine Learning models classification results



The diagram 17 shows the summary of the classification performance results of traditional baseline models and machine learning models. According to the chart, amongst the traditional baseline methods, Bollinger Bands baseline model, with an overall accuracy level of 57.53, is outperforming RSI and MACD baseline models whose classification accuracy results are 50.41% and 44.06%, respectively. On the contrary, the Hybrid model (Random Forest and Gradient Boosting) achieved a better overall classification accuracy level compared to all traditional baseline models and all other machine learning models applied in this study.

4.4. Regression results and discussion

Diagram 18. Regression results of Random Forest, XGBoost, Gradient Boosting and Hybrid models.

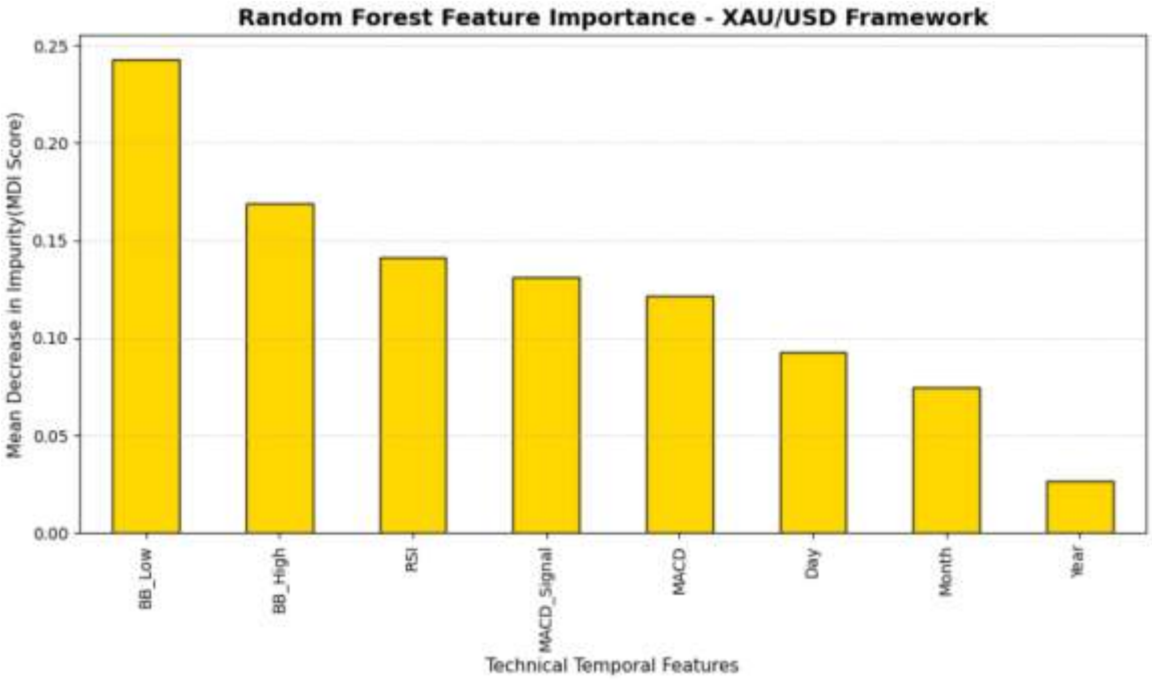
```
RANDOM FOREST REGRESSION
RMSE: 167.10062723039383
MAE: 152.64636539704145
R^2: -6.143317831065327

XGBOOST REGRESSION
RMSE: 171.62015204388038
MAE: 150.10963327687935
R^2: -6.534950063859915

GRADIENT BOOSTING REGRESSION
RMSE: 164.88697282715464
MAE: 145.4365099171692
R^2: -5.9553101918386595

HYBRID REGRESSION (ENSEMBLE AVERAGE)
RMSE: 167.03002527163602
MAE: 149.34285900169937
R^2: -6.137282836509735
```

According to the diagram 18, regression-based forecasting evaluation generated similar comparative findings among the machine learning models. The Gradient Boosting regression model achieved lower forecasting errors compared to the Random Forest, XGBoost and the Hybrid regression models, meaning that the Gradient Boosting has relatively stronger numerical forecasting capability in contrast to other models used in this study. To be more specific, the Gradient Boosting regression model achieved an RMSE value of around 164.89 and an MAE value of about 145.44. In contrast, the Random Forest, XGBoost and Hybrid machine learning regression models produced larger forecasting errors, with RMSE and MAE values of around 167.10 and 152.65 for Random Forest, 171.62 and 150.11 for XGBoost and 167.03 and 149.34 for Hybrid regression model. In other words, the forecasted gold prices of the Gradient Boosting model deviated less from actual market prices compared to other machine learning results. This implies that the Gradient Boosting model captured nonlinear relations between technical indicators and gold price movements more effectively than other machine learning models under the current forecasting environment. However, despite the comparatively lower forecasting errors, all regression models produced negative R^2 values. The Gradient Boosting regression model generated an R^2 value of approximately -5.96, while Random Forest, XGBoost and Hybrid models produced even lower values of approximately –6.14, -6.53 and -6.14, respectively. Negative R^2 values mean that all machine learning models performed worse than a simple baseline prediction using the mean of the dependent variable. From an analytical perspective, this outcome demonstrates that exact numerical forecasting of gold prices remains extremely difficult when relying primarily on technical indicators and temporal variables. The weak regression performance highlights the stochastic and highly dynamic nature of XAU/USD markets. Gold prices are heavily influenced by external macroeconomic and geopolitical factors including Federal Reserve policy decisions, inflation announcements, recession expectations, global conflicts, currency fluctuations, and safe-haven investment demand. Since these macroeconomic variables were not incorporated into the present framework, the machine learning models faced structural limitations in explaining broader market behavior.

The results therefore imply that technical indicators may possess stronger capability for capturing directional tendencies rather than exact future price levels. In practical financial applications, technical indicators are often more effective for identifying momentum shifts and trading signals than for producing highly precise numerical forecasts. Nevertheless, the superior RMSE and MAE performance of Gradient Boosting compared to Random Forest, XGBoost and the Hybrid models suggests that Gradient Boosting extracted more meaningful nonlinear information from the technical indicator dataset.

4.5 Feature Importance Analysis and Discussion

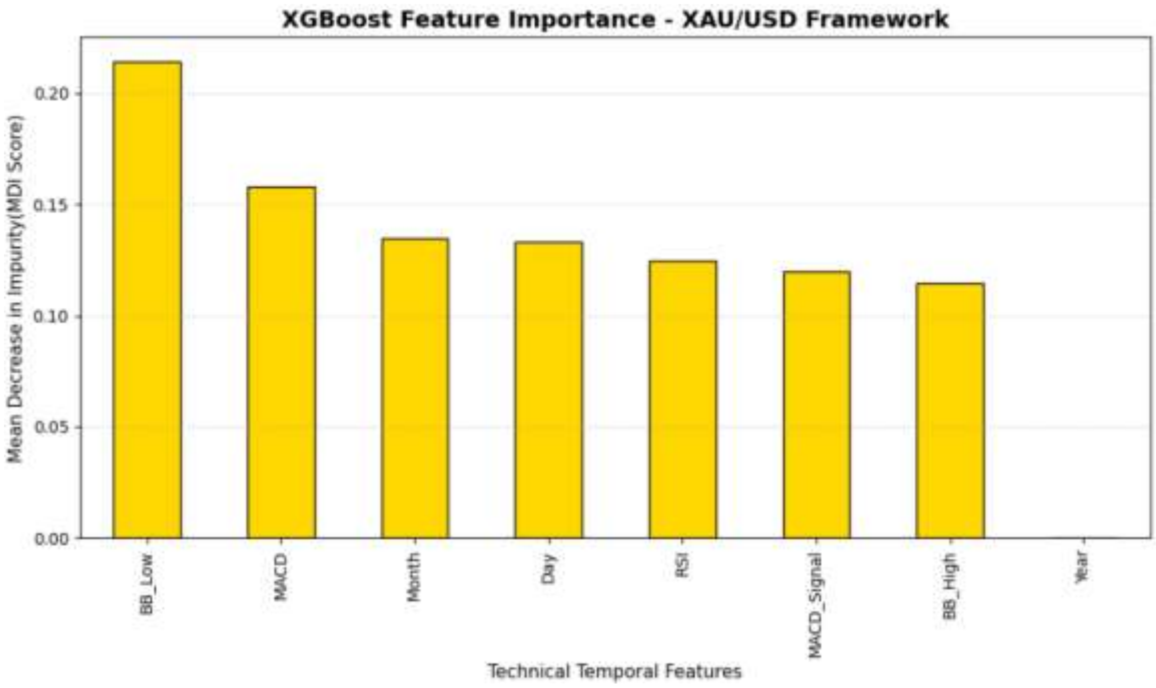
Diagram 19. Random Forest Feature Importance



The diagram 19 shows the feature importance scores generated by the Random Forest model. The feature importance is evaluated by the Mean Decrease in Impurity (MDI), which indicates the contribution of each predictor variable to reducing prediction error across all decision trees within the Random Forest ensemble. According to the results, both Bollinger Band Low (BB_Low) and Bollinger Band High (BB_High) are two most influential predictors, with importance scores of around 0.24 and 0.17, respectively, together accounting for over 40% of the model's predictive power. This highlights the significance of Bollinger Bands in recognizing market volatility and potential price reversals. This finding also means that Bollinger Bands indicator provides substantial information about future gold price movements, the model appears to rely heavily on this indicator when generating forecasts. Moreover, among the remaining technical indicators, RSI demonstrates relatively strong importance (around 0.14), followed by MACD Signal (around 0.13) and MACD (around 0.12). This finding implies that momentum-based indicators contribute meaningfully to the forecasting process, even though their impact is lower than that of the Bollinger Band variables. Therefore, the Random Forest model relies on both volatility-based and momentum-based information when predicting gold prices. Furthermore, calendar features, including Day, Month and Year have comparatively lower importance. Day, and month account for 0.10 and 0.08, respectively. This finding indicates that short-term seasonal effects may provide some predictive value but are less influential than the technical indicators.

The year variable accounts for about 0.03, indicating that long-term calendar effects have limited predictive power within the forecasting framework.

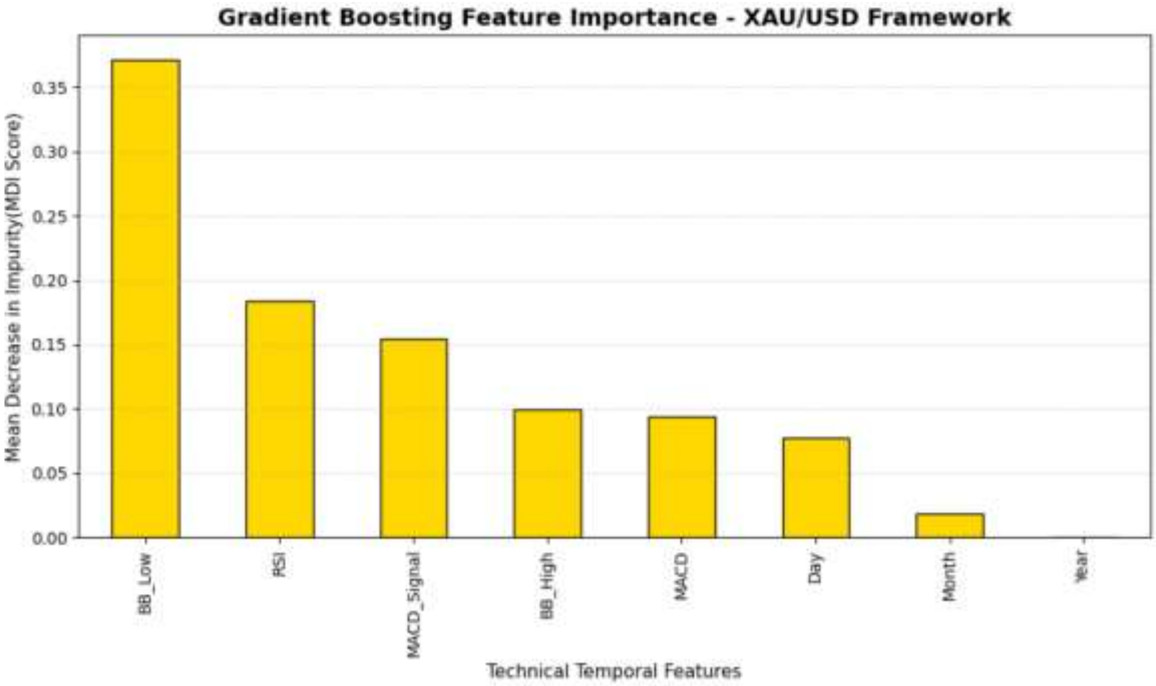
Diagram 20. XGBoost Feature Importance



The Diagram 20 illustrates the feature importance scores generated by the XGBoost model.

According to the results, Bollinger Band Low (BB_Low) is the most influential predictor, with an importance score of approximately 0.22, meaning that the XGBoost heavily relies on lower Bollinger Band in forecasting gold prices. It also means that volatility based indicators are significantly important in forecasting gold price movements. The second most influential indicator is MACD, accounting for an importance score of about 0.16, indicating that the trend-following and momentum information is vital in the XGBoost forecasting process. In other words, it suggests that changes in market momentum significantly influence gold price movements and provide valuable predictive information for the model. Furthermore, the variables Month and Day have moderate role in gold price forecasting, each accounting for importance scores of approximately 0.13. The higher importance level of these temporal features compared with the Random Forest model indicates that the XGBoost model captures seasonal and calendar-related patterns more effectively. It might be implied that in some cases monthly or daily market behaviors may contribute to fluctuations in gold price movements. On the contrary, the RSI, MACD Signal, and BB_High variables account for similar levels of importance, ranging between 0.11 and 0.12. They some contribution to the forecasting process, but their influence is lower than that of BB_Low and MACD. It might also suggest that the XGBoost model benefits from integrating multiple sources of market information, including volatility, momentum, and trend confirmation signals. Finally, the Year variable has no importance, with a score close to zero. This finding suggests that long-term calendar effects provide little or no predictive value for the XGBoost model and do not significantly contribute to forecasting gold price movements.

Diagram 21. Gradient Boosting Feature Importance



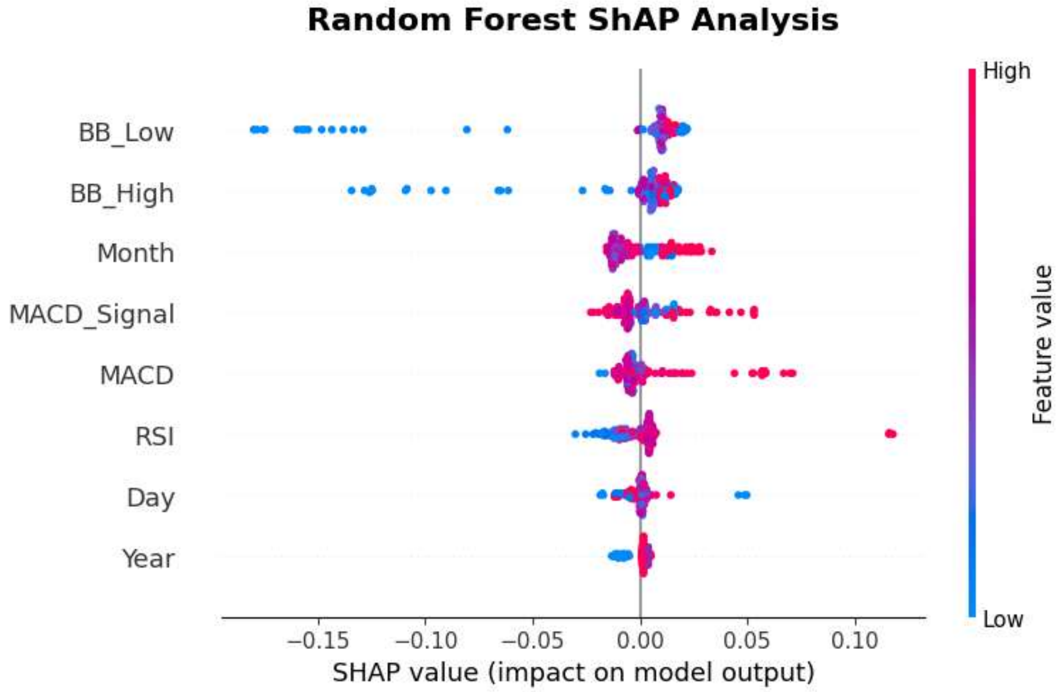
The Diagram 21 shows the feature importance scores generated by the Gradient Boosting model. The results show that Bollinger Band Low (BB_Low) accounts for an importance score of around 0.37, making it the most influential predictor. This value is substantially higher than that of all other variables, indicating that the Gradient Boosting model relies heavily on the lower Bollinger Band when forecasting gold prices. This also suggests that price movements near the lower volatility boundary contain significant predictive information about future gold price behavior. Moreover, the second most important feature is RSI, with an importance score of approximately 0.18, meaning that market momentum and potential overbought or oversold conditions play an important role in the forecasting process. Furthermore, the MACD Signal accounts for an importance score of approximately 0.15, highlighting the relevance of trend-confirmation information. This result might also suggest that the Gradient Boosting model benefits from identifying changes in market momentum and trend direction before major price movements happen. Besides that, the upper Bollinger Band (BB_High) and MACD have equal importance levels, both accounting for importance scores of around 0.10. Even though their contribution is considerably lower compared to that of BB_Low, the Gradient Boosting framework uses a combination of volatility and trend-based indicators to improve forecasting accuracy. Furthermore, regarding the calendar features, the variable Day has a small contribution, with an importance score of approximately 0.09, indicating the presence of limited short-term calendar effects. In contrast, the variable Month shows very low importance, around 0.03, while the variable Year almost has no contribution to the model.

4.6 SHAP Analysis and Discussion

In order to improve the interpretability of the machine learning models, SHAP (SHapley Additive exPlanations) analysis was conducted.

SHAP values quantify the contribution of each predictor variable to the model's output, allowing both the magnitude and direction of each feature's influence to be examined. Now let us discuss the SHAP analysis results for each machine learning model used in this study.

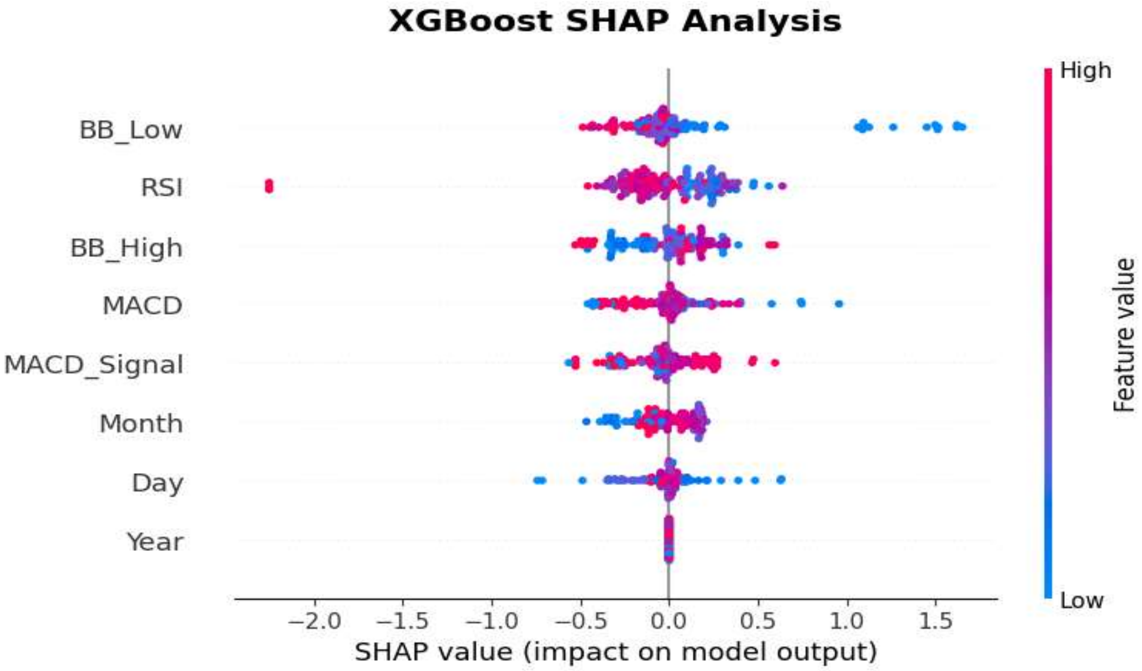
Diagram 22. Random Forest SHAP analysis



The diagram 22 shows the Random Forest SHAP analysis, ranking the variables based on their average absolute contribution to the model predictions. According to the results, Bollinger Bands Low (BB_Low) and Bollinger Bands High (BB_High) are the most influential predictors, accounting for the largest SHAP value dispersion, suggesting that they have the strongest impact on the model's predictions. Their SHAP values range approximately from -0.17 to +0.03, substantially larger than those of the remaining predictors. Since Bollinger Bands capture price volatility and relative price positioning within a trading range, the results suggest that volatility-related information is highly informative for the target variable. Moreover, the observations with lower feature values (blue points) tend to produce more negative SHAP values, implying that lower Bollinger Band levels generally decrease the predicted outcome. Conversely, higher values contribute positively, although the positive effect is comparatively smaller. Furthermore, the variable Month is the most important calendar feature in this model. The distribution of SHAP values indicates a clear relationship between specific months and prediction outcomes. In other words, the higher month values (red points) are generally associated with positive SHAP values, whereas the lower month values contribute negatively, indicating that seasonality effects are present within the dataset. On the other hand, the variables Day and Year account for relatively small SHAP ranges centered around zero. Their limited contribution implies that daily and yearly temporal fluctuations provide little additional predictive information beyond what is already captured by the technical indicators. Besides that, both MACD and MACD_Signal demonstrate moderate predictive importance. Higher values of these indicators are predominantly associated with positive SHAP values, indicating that stronger bullish momentum contributes positively to the model's output.

The consistency of this relationship aligns with financial theory, as MACD-based indicators are widely used to identify trend strength and momentum shifts. The Random Forest model appears to utilize these momentum signals to distinguish between favorable and unfavorable prediction outcomes. In contrast, the Relative Strength Index (RSI) has less contribution than the Bollinger Band and MACD variables but still provides useful predictive information. Most observations cluster around zero SHAP values, indicating a generally moderate effect. However, several observations exhibit relatively large positive SHAP values, suggesting that extreme RSI conditions may significantly influence predictions in specific cases, which can imply that RSI plays a supplementary role, enhancing prediction accuracy under particular market conditions rather than acting as a primary determinant of model behavior.

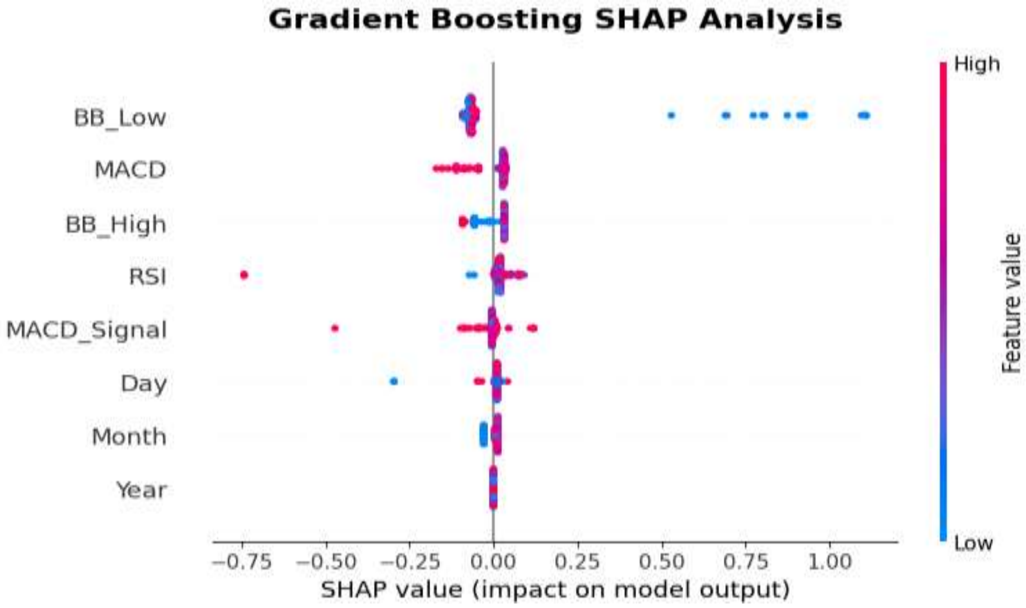
Diagram 23. XGBoost SHAP Analysis



The diagram 23 shows the XGBoost SHAP analysis, ranking the variables based on their average absolute contribution to the model predictions. So, the results indicate that Bollinger Bands Low (BB_Low) is the most influential predictor, accounting for the largest SHAP value dispersion, indicating that it has the strongest influence on the XGBoost model's predictions. The SHAP values range approximately from -0.5 to +1.7, which is substantially larger than the ranges observed for the remaining variables. The presence of several observations with exceptionally large positive SHAP values suggests that BB_Low can strongly increase the model output under specific market conditions. At the same time, many observations remain clustered around zero, indicating that the effect of BB_Low is not uniform across all instances. This pattern suggests that the XGBoost model has learned complex non-linear relationships involving Bollinger Band information rather than a simple monotonic effect. Moreover, the XGBoost model identifies RSI as the second most important predictor.

The wide dispersion of SHAP values suggests that RSI contributes substantially to prediction outcomes. Most observations are concentrated near zero, but a notable negative outlier with a SHAP value below -2.0 indicates that RSI can have an exceptionally strong influence under certain conditions, meaning that the model is particularly sensitive to extreme RSI values, potentially capturing overbought or oversold market states that significantly affect future price movements. This can mean that the XGBoost model is able to extract more predictive information from momentum-related signals than the Random Forest model. Furthermore, the Bollinger Bands High (BB_High) is the third most contributing variable to model predictions. The distribution of SHAP values on both sides of zero indicates that BB_High can either increase or decrease predictions depending on the market context, suggesting that Bollinger Band indicators provide critical information about market volatility and price positioning, which the XGBoost model utilizes extensively. Moreover, both MACD and MACD_Signal have moderate predictive importance, whose SHAP values are distributed around zero, with several observations extending toward positive values. This might indicate that momentum signals contribute to prediction formation but are not as dominant as Bollinger Band indicators or RSI. The moderate influence of MACD-based variables is consistent with their role as trend-following indicators. The XGBoost model appears to incorporate these momentum measures as complementary signals that improve predictive accuracy when combined with volatility-related features. Regarding the calendar features, the variable Month has the greatest predictive relevance. Although its SHAP value range is considerably smaller than that of the technical indicators, the variable contributes more substantially than both Day and Year. The SHAP value of the variable Month near zero indicates that seasonal effects exist but are relatively weak compared with technical indicators. Therefore, while certain months may influence predictions, the XGBoost model relies far more heavily on market-derived features than on calendar-based information. Furthermore, both variables Day and Year exhibit SHAP values that remain tightly clustered around zero. Their limited dispersion suggests that these variables provide minimal additional information and have little influence on prediction outcomes.

Diagram 24. Gradient Boosting SHAP Analysis



The diagram 24 shows the Gradient Boosting SHAP analysis, ranking the variables based on their average absolute contribution to the model predictions. According to the results, the

Bollinger Bands Low (BB_Low) is the most influential predictor, accounting for the largest SHAP value dispersion among all predictors, suggesting that it has the strongest impact on the Gradient Boosting model's predictions. The Bollinger Bands Low variable has exceptionally large positive SHAP values exceeding +1.0, indicating that BB_Low can substantially increase the model output under specific market conditions. At the same time, most observations remain concentrated near zero, implying that the impact of BB_Low is not consistent across all instances, suggesting that the Gradient Boosting model relies heavily on BB_Low only in particular market scenarios, while its effect remains limited for many observations. Furthermore, among the momentum indicators, MACD is the second most influential predictor in the Gradient Boosting model. The distribution of SHAP values suggests that MACD contributes meaningfully to prediction outcomes, with observations appearing on both sides of the zero line. The prominence of MACD indicates that trend strength and momentum information are highly relevant for the model. Compared with the Random Forest model, where Bollinger Band indicators dominated the prediction process, the Gradient Boosting model appears to place relatively greater emphasis on momentum signals captured by MACD. Besides that, the Bollinger Bands High (BB_High) is the third most contributing variable to the prediction process. Even though its SHAP range is smaller than that of BB_Low, the variable demonstrates both positive and negative contributions to the model output, indicating that information contained within Bollinger Band indicators plays a central role in the model's decision-making process, highlighting the importance of volatility-related market signals. In contrast, the MACD_Signal variable also contributes to the prediction process, although its influence is weaker than that of MACD. Most observations are concentrated near zero, indicating that MACD_Signal acts primarily as a supporting predictor rather than a dominant driver of model behavior. Nevertheless, several observations display notable negative SHAP values, suggesting that the signal line can significantly influence predictions under certain market conditions. Furthermore, the Relative Strength Index (RSI) demonstrates moderate predictive importance within the Gradient Boosting model. Most SHAP values are clustered around zero, indicating that RSI generally exerts a limited influence on predictions. However, the presence of a substantial negative outlier suggests that RSI can become highly influential in specific observations. This might suggest that RSI primarily serves as a conditional predictor whose importance increases during extreme market situations. Consequently, RSI appears to complement Bollinger Band and MACD indicators rather than acting as a primary determinant of prediction outcomes. Regarding the calendar features, the temporal variables (Day, Month, and Year) exhibit relatively small SHAP value ranges centered around zero, whose limited dispersion suggests that they provide little additional predictive information compared with the technical indicators. The variable Day shows slightly greater variation than Month and Year even though its overall contribution remains modest. The near-zero SHAP values associated with the variable Year indicate that long-term calendar effects have minimal influence on the model's predictions. Moreover, the variable Month contributes only marginally, suggesting that seasonal patterns are considerably less important than market-derived technical indicators.

5. Legal and ethical considerations

In terms of legal considerations, we took deliberate steps to ensure that all data used in this study were sourced from public financial platforms, such as widely recognized market websites. We carefully verified that no proprietary or copyrighted data were used without permission, ensuring full compliance with intellectual property laws. Ethically, we were diligent in our transparency, meaning that every data source, preprocessing step, and analytical method was fully disclosed. We avoided any selective reporting, and all results were openly shared so that others could replicate or scrutinize them. Professionally, we ensured that the study was conducted independently, free from any financial institution's influence, so there were no conflicts of interest. We also adhered to professional best practices in financial modeling, following established academic guidelines. However, we remain aware that financial data governance is evolving, and we must stay attuned to new regulations or data privacy standards as they emerge, ensuring that future research continues to respect legal and ethical boundaries.

6. Limitations of the study

This study has several limitations that must be acknowledged. First, the model is based exclusively on historical price data and technical indicators, which means it does not incorporate macroeconomic factors. For example, we did not include variables such as inflation rates, real interest rates, or global geopolitical events, all of which can significantly drive gold price movements. This omission means the model may miss key drivers of price changes, limiting its ability to fully explain market dynamics. Second, the dataset spans only a specific time period, and market conditions within that window may not reflect future or broader trends. As a result, the model's generalizability is constrained, meaning that it may perform well on this data but fail under different market regimes. Third, despite using cross-validation, the hybrid model still faces risks of overfitting. Given the imbalance in the class distribution, where one class was much more frequent, the model may have learned patterns that reflect the sample rather than the underlying market behavior. Finally, the model is static, meaning it does not adapt to changing market conditions or incorporate real-time data, which limits its practical deployment in dynamic trading environments. Thus, while the hybrid model advances forecasting beyond traditional indicators, these limitations highlight the need for future research to integrate macroeconomic factors, expand the timeframe, and adopt adaptive modeling approaches to improve robustness and applicability.

7. Conclusion

This study developed and evaluated a hybrid forecasting framework for gold price prediction by integrating technical indicators with ensemble machine learning algorithms. Using historical XAU/USD data, technical indicators including RSI, MACD, MACD Signal, and Bollinger Bands were employed to forecast gold price movements through both classification and regression approaches. The performance of traditional technical indicator models was compared with Random Forest, XGBoost, Gradient Boosting, and several hybrid machine learning models.

The findings demonstrated that machine learning models generally outperformed traditional technical analysis methods. Among the baseline models, Bollinger Bands achieved the highest classification accuracy, while the proposed hybrid model combining Random Forest and Gradient Boosting produced the best overall forecasting performance with an accuracy of 60.14%.

The results of Time Series Cross-Validation further confirmed the relative stability of the machine learning models across different time periods. Feature Importance and SHAP analyses revealed that Bollinger Band variables, particularly BB_Low, were the most influential predictors, highlighting the importance of volatility-related information in forecasting gold price movements.

However, the regression results showed relatively weak performance, with all models producing negative R^2 values despite Gradient Boosting achieving the lowest forecasting errors.

This suggests that technical indicators are more effective for predicting the direction of future price movements than for forecasting exact gold price levels. Overall, the study concludes that integrating technical indicators with ensemble machine learning techniques can improve gold price directional forecasting and provides a more effective framework than traditional standalone technical analysis methods.

8. Recommendations

Several recommendations can be proposed for future research. First, future studies should incorporate macroeconomic variables such as inflation rates, interest rates, exchange rates, central bank policies, and geopolitical risk indicators, as these factors significantly influence gold price behavior and may improve forecasting accuracy. Second, expanding the dataset to include a longer historical period could improve model generalizability and enable the framework to capture different market regimes. Third, future research may explore advanced deep learning techniques such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Transformer-based models to determine whether they can outperform traditional ensemble machine learning methods.

Fourth, real-time and adaptive forecasting frameworks should be investigated to allow models to adjust dynamically to changing market conditions. Finally, future studies should consider combining technical indicators, macroeconomic variables, and market sentiment indicators within a unified hybrid framework to further improve gold price forecasting performance and practical applicability.

References

1. Amini, M. and Kalantari, M. (2024) ‘Gold price prediction using deep learning models: A hybrid CNN–BiLSTM approach’, Expert Systems with Applications, 237, 121615.
2. Baur, D.G. and Lucey, B.M. (2010) ‘Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold’, Financial Review, 45(2), pp. 217–229.
3. Baur, D.G. and McDermott, T.K. (2010) ‘Is gold a safe haven? International evidence’, Journal of Banking and Finance, 34(8), pp. 1886–1898.
4. Beckmann, J., Berger, T. and Czudaj, R. (2015) ‘Does gold act as a hedge or a safe haven? New evidence from time-varying coefficient models’, Economic Modelling, 48, pp. 16–24.
5. Bollinger, J. (2001) Bollinger on Bollinger Bands. New York: McGraw-Hill
6. Breiman, L. (2001) ‘Random forests’, Machine Learning, 45(1), pp. 5–32.
7. Brock, W., Lakonishok, J. and LeBaron, B. (1992) ‘Simple technical trading rules and the stochastic properties of stock returns’, Journal of Finance, 47(5), pp. 1731–1764.
8. Chen, T. and Guestrin, C. (2016) ‘XGBoost: A scalable tree boosting system’, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

9. Chong, T.T.L. and Ng, W.K. (2008) ‘Technical analysis and the London Stock Exchange’, Applied Economics Letters, 15(14), pp. 1111–1114.
10. Fernando, J. (2025) ‘Relative Strength Index (RSI): What it is, how it works, and formula’. Investopedia. Available at: https://www.investopedia.com/terms/r/rsi.asp
11. Fidelity Investments (2026) ‘Bollinger Bands’. Available at: https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/bollinger-bands
12. Fidelity Investments (2026) ‘Relative Strength Index (RSI)’. Available at: https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/rsi
13. Fidelity Investments (2026) ‘The Moving Average Convergence/Divergence indicator’. Available at: https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/macd
14. Fischer, T. and Krauss, C. (2018) ‘Deep learning with long short-term memory networks for financial market predictions’, European Journal of Operational Research, 270(2), pp. 654–669.
15. Friedman, J.H. (2001) ‘Greedy function approximation: A gradient boosting machine’, Annals of Statistics, 29(5), pp. 1189–1232.
16. Ghosh, D., Levin, E.J., Macmillan, P. and Wright, R.E. (2004) ‘Gold as an inflation hedge?’, Studies in Economics and Finance, 22(1), pp. 1–25.
17. Gu, S., Kelly, B. and Xiu, D. (2020) ‘Empirical asset pricing via machine learning’, Review of Financial Studies, 33(5), pp. 2223–2273.
18. Gupta, R. (2025) ‘Disadvantages of limitations of technical analysis’. Available at: https://www.stockgro.club/blogs/trading/limitations-of-technical-analysis
19. Hoseinzade, E. and Haratizadeh, S. (2019) ‘CNN-based stock market prediction using technical indicators’, Expert Systems with Applications, 136, pp. 46–57.
20. Jiang, W. (2021) ‘Applications of machine learning in stock market prediction: A survey’, Journal of Physics: Conference Series, 1861(1), pp. 1–8.
21. Kim, K.J. (2003) ‘Financial time series forecasting using support vector machines’, Neurocomputing, 55(1–2), pp. 307–319.
22. Konur, O., Göçken, M. and Dosdoğru, A.T. (2024) ‘The role of technical indicators in machine learning-based financial forecasting’, Journal of Open Innovation: Technology, Market, and Complexity, 10(2), 100215.
23. Krauss, C., Do, X.A. and Huck, N. (2017) ‘Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500’, European Journal of Operational Research, 259(2), pp. 689–702.
24. Krollner, B., Vanstone, B. and Finnie, G. (2010) ‘Financial time series forecasting with machine learning techniques: A survey’, European Symposium on Artificial Neural Networks, pp. 389–394.
25. Lo, A.W., Mamaysky, H. and Wang, J. (2000) ‘Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation’, Journal of Finance, 55(4), pp. 1705–1765.

26. Ma, Z., Wang, X. and Hao, Y. (2023) ‘Development and application of a hybrid forecasting framework based on improved extreme learning machine for enterprise financing risk’, Expert Systems with Applications, 215, 119373.
27. Natekin, A. and Knoll, A. (2013) ‘Gradient boosting machines, a tutorial’, Frontiers in Neurorobotics, 7(21), pp. 1–21
28. Neely, C.J., Rapach, D.E., Tu, J. and Zhou, G. (2014) ‘Forecasting the equity risk premium: The role of technical indicators’, Management Science, 60(7), pp. 1772–1791
29. Nguyen, T.H. (2024) ‘Forecasting gold price using technical analysis indicators’, International Journal of Financial Studies, 12(1), 15.
30. Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015) ‘Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques’, Expert Systems with Applications, 42(1), pp. 259–268.
31. Sezer, O.B., Gudelek, M.U. and Ozbayoglu, A.M. (2020) ‘Financial time series forecasting with deep learning: A systematic literature review’, Applied Soft Computing, 90, 106181.
32. Vrtagic, S. and Mourched, M. (2025) ‘Hybrid machine learning models for commodity price forecasting: Evidence from gold markets’, International Journal of Computational Methods and Experimental Measurements, 13(1), pp. 45–60
33. Zhang, X., Aggarwal, V. and Qi, G. (2017) ‘Stock price prediction via discovering multi-frequency trading patterns’, Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2141–2149.